University of Warsaw
Faculty of Economic Sciences

Working Papers
No. 14/2011 (54)

Michał Krawczyk

Framing in the field. A simple experiment on the reflection effect

Warsaw 2011

Framing in the field.
A simple experiment on the reflection effect

Michał Krawczyk
University of Warsaw
Faculty of Economic Sciences
e-mail: mkrawczyk@wne.uw.edu.pl

Abstract

This study makes use of an unusual opportunity to manipulate framing of a simple decision under uncertainty: whether or not to answer an exam question when unsure which answer is correct and a missing response is scored higher than an incorrect one. Two treatments were compared in a natural field experiment: one in which the decision was framed in terms of losses, and the other – in terms of gains. Some alternative theories of decision making under risk, notably prospect theory, propose that individuals display reflection effect, i.e. tend to be more risk-seeking in losses than gains. No such evidence was found: subjects were generally risk-averse and this disposition was not affected by treatment.

Keywords:
framing, reflection effect, field experiments

JEL:
C93, D81

1 Introduction

The leading alternative theory of decision making under risk—the cumulative prospect theory (Tversky and Kahneman 1992), proposes that most individuals are generally reluctant to accept sure losses. They may be willing to assume additional risk if it creates a chance to break even. In other words, individuals are risk-averse in gains and risk-seeking in losses, at least for events of moderate probability (*reflection effect*). One well known example of this phenomenon is the disposition effect (Odean 1998) – the tendency to hold on to losing stocks, in the hope that the price goes back up again (while winners would often be sold quickly). Another interesting case is the withholding phenomenon – the tendency of taxpayers to show more compliance when they are owed a refund than when they owe additional tax (Clotfelter 1983). However, these and most of other empirical cases supporting the reflection effect are fairly complex phenomena and alternative explanations are often possible.

Perhaps the most substantial bulk of evidence in favor of the reflection effect comes from laboratory experiments on framing. In this approach, largely inspired by the "Asian disease" problem of lives lost vs. lives saved by Tversky and Kahneman (1981), a hypothetical gamble is either presented in terms of losses or gains. A meta-analysis of this literature (Kuhberger 1998) confirms a moderate effect of greater risk seeking in losses. Experiments with real incentives are much less common (but see e.g. Kuhberger, Schulte-Mecklenbeck, and Perner (2002)), chiefly because inducing a feeling of loss in the lab is not trivial.

To date, there have been very few *field* experiments aimed at identifying and possibly measuring risk-seeking in losses. The main methodological problem is that it is difficult to find situations in which a sizable population is involved in simple case of decision making under risk with non-trivial stakes and potential losses that can be actively manipulated by the researcher, observed and registered without the participants' knowledge and consent. In this study I make use of a unique opportunity to induce perception of a possible loss in a simple situation of decision making under uncertainty. The findings can be summarized succinctly: subjects are just as risk-averse in losses as they are in gains and it seems to be difficult to ascribe this null result to any of the "usual suspects" – sample size, stakes, confusion, boredom, ceiling effect, etc. In view of the predominantly positive findings in previous studies, it is an intriguing result that calls for future research. It may suggest that the reflection effect is not as robust as it is sometimes believed to be.

2 Design

The study was performed during an examination at the University of Warsaw - the final exam in Microeconomics at the Faculty of Economic Sciences. The exam involved 30 multiple-choice questions (with one correct and three to four incorrect answers each). Unbeknownst to the students, exam instructions came in two types, as far as the number of points per item was concerned: the Loss Treatment (LT) and the Gain Treatment (GT), see Table 1.

Table 1: *Numbers of points awarded per answer*

treatment	correct	missing	incorrect
LT	0	-2	-3
GT	3	1	0

Note that one could never lose a point in GT nor gain a point in LT. Whatever answer was given to any question, a student would be awarded three points more in the GT than in the LT.[1] As a result, any combination of answers would automatically bring 90 points more in the gain than in the loss condition. Accordingly, students in the loss condition would start with the score of 90 points.

With regards to the information provided to the students, before the exam they had only been told that leaving a question unanswered would yield more points than giving an incorrect answer. Assignment to exam rooms and, as a consequence, to the two treatments was random. At the start of the exam, the students were given the accurate description of the scoring rules relevant for their treatment only (the translation of the exam copies are available from the author). They were also told that most likely approx. 45 points would be necessary to pass.

After the exam, the students were asked to predict their score. They were told that one of those who get it right (or else are closest) would obtain a monetary reward of 250 PLN (approx. 62 euro).

One hundred and fifty three students took part in the exam and the experiment, and submitted their predictions (one student skipped all the exam questions as well as the prediction and was removed from the database). The data collected involved the gender of the student, his or her grade obtained in

[1] Bosch-Domènech and Silvestre (2006) call it "translation"

the tutorials, name of the teaching assistant, the number of correct, incorrect, and missing answers and finally the predicted score.

3 Predictions

Choosing the best answer is obviously a task involving a subject-specific skill and knowledge. However, whether to actually choose it or rather leave the question unanswered is a decision under uncertainty.

Crucially, for any level of certainty about one's ability to pick the right answer, answering the question is a risky option, while giving no answer guarantees a fixed payoff. Prospect theory proposes that for a moderate probability of the highest of two outcomes, most individuals will be risk-averse in the domain of gains but risk-seeking in the domain of losses. This pattern results from the curvature of the value function, being concave for gains and convex for losses and the shape of the probability weighting function.

More precisely, let us denote the minimum individually perceived probability of selecting the right answer that would induce an individual to take the risk under GT as p. With standard prospect-theoretic notation we have $w^+(p)v(3) = v(1)$. Suppose for a moment that probability weighting is linear, then we have that $pv(3) = v(1)$ so if v is concave for gains, it implies that $p > 1/3$. We now ask what this person would do under LT. Generally, she will take the risk if $w^-(1-p)v(-3) > v(-2)$ which under linear probability weighting simplifies to $(1-p)v(-3) > v(-2)$. If v is convex for losses and $1 - p < 2/3$, this will indeed be satisfied. Thus convex/concave value function with correct probability weighting implies that the minimum level of certainty required to answer the question is lower for LT than GT. Consequently, more risk will be taken (less questions skipped) under LT. However, studies on Cumulative Prospect Theory further claim that a typical subject overweights low probabilities (say, below .3) and underweights moderate and high probabilities. If p is moderate for most subjects, perhaps between .3 and .7, as we will later show, then both "extreme" events (obtaining 3 or -3) are underweighted, which further strengthens the risk-loving tendency under LT than GT.

There are two important caveats to this prediction. First, obviously, it is not clear whether endowing subjects with 90 points up-front and then slowly taking them away truly induces a feeling of a loss. However, the "windfall" nature of such endowment appears, if anything, to induce even more risk-seeking ("house money effect", Thaler and Johnson (1990)).

Second, participants' behavior may be entirely driven by willingness to obtain a possibly highest grade, rather than maximizing the number of points.

For example, if a student contemplating whether or not to try to answer any given question knows that of the remaining 29 problems he has selected the correct answer to 15 questions and given no answer to 14 questions (obtaining 59 points under GT) and furthermore that the next grade threshold is 60, it seems an obvious choice to skip the remaining question. This is true under both treatments. However, this reasoning rests on strong assumptions.

First, that participants are indifferent to various point scores as long as they obtain the same grade. Still, students in the elite academic program under scrutiny are typically highly motivated; scoring 34 points out of 90 is less unpleasant and humiliating than ending up with just 15, although both would result in the same grade. In this sense, the grade is not all that matters. Furthermore, losing by a small margin only, students may generally expect more leniency on the part of the lecturer (e.g. his or her willingness to grant an additional chance to pass).

The second implicit assumption is that students are able to perfectly predict their score and thus tell when they are, say, just below the threshold. As we will see when analyzing the prediction data, this was not the case.

More generally, this kind of concern applies equally to monetary rewards: the value of money is in the goods and services that it can buy. It seems that in our case points could play the role of a "prime reinforcer," just as we generally believe money does in standard lab experiments.

As for other effects that we may expect, a great many studies show that males are generally more willing to take risk than females (Byrnes, Miller, and Schafer 1999), especially when it is based on assessment of own ability. We therefore predict less unanswered questions in male subjects.

To provide an additional measure of risk aversion, one of the questions was designed to be extremely difficult (and explicitly labeled as such), involving a tricky integral that students were very unlikely to be able to solve under strict time constraints. Three of the four answers were numerically very close to one another (and hence equally likely to be correct, unless one was able to actually solve the problem), while the fourth one was absurdly high (equal to 100^{100}). Thus from the perspective of a subject in GT, answering this question was tantamount to taking a lottery (3,1/3), while skipping it resulted in a payoff of 1 for sure, i.e. (1,1). Similarly, those in the LT had a choice between $(-3, 2/3)$ and $(-2, 1)$. Risk-seeking individuals should thus answer the question, while risk-averters should skip it.

4 Results

Overall, students left 7.58 questions unanswered. This number varied substantially between subjects, the standard deviation was 4.47. Only two students tried to answer all the questions.

The treatment effect turns out to be small and insignificant. Indeed the mean number of missing answers among the 79 subjects in the Gain Treatment was 7.76, while the mean of 74 observations in the LT was 7.09. The difference is not significant ($p = 0.31$ in a simple two-sided t-test, $p = 0.45$ in a nonparametric Mann-Whitney test).

Regarding the malicious unsolvable integral, only eight out of 153 subjects (five in LT, three in GT) tried to guess the answer. None of them chose the ostensibly high 100^{100}, but only three actually picked the right answer, in line with our assumption that the questions represented a $(3, 1/3)$ gamble. Therefore we conclude that subjects were generally risk-averse.[2] Figure 1 suggests that these 8 subjects did not seem to be systematically different from others in terms of their predictions or actual scores. There is thus no evidence of risk posture being affected by own position with respect to thresholds. As a matter of fact, this general unwillingness to take a fair bet even among individuals who (correctly) expect to be below the passing threshold and should thus like to increase the variance of their score is remarkable. It may perhaps be best explained in terms of ambiguity aversion, especially if students imagine they could subsequently be asked to justify their choice (Trautmann, Vieider, and Wakker 2008).

Regarding subjects' predictions, Figure 1 shows that they had a general idea about their score but they were not able to predict it very precisely and were somewhat overly confident or optimistic (there are more observations above the 45° line than below). It also reveals that treatment had no impact on prediction optimism or accuracy.

It may also be instructive to investigate the individually perceived success rates for the questions the subjects did answer that are implied by their predictions. For example, if a student answered 20 questions (thus skipped ten) and then predicted a score of 43, it means, she expected to get 33 points for the questions she did answer (thus have 11 correct answers), so her implied success rate will be 11/20. Figure 2 shows the distributions of predictions-implied success rates for the two treatments.

Again, not surprisingly, we do not see any treatment effect here either. Secondly, note that there are very few (approx. 5%) values outside of the

[2] This also validates our claim that the minimum probability of success necessary to take the risk is above .3 for vast majority of the subjects.

Figure 1: Predicted and actual score

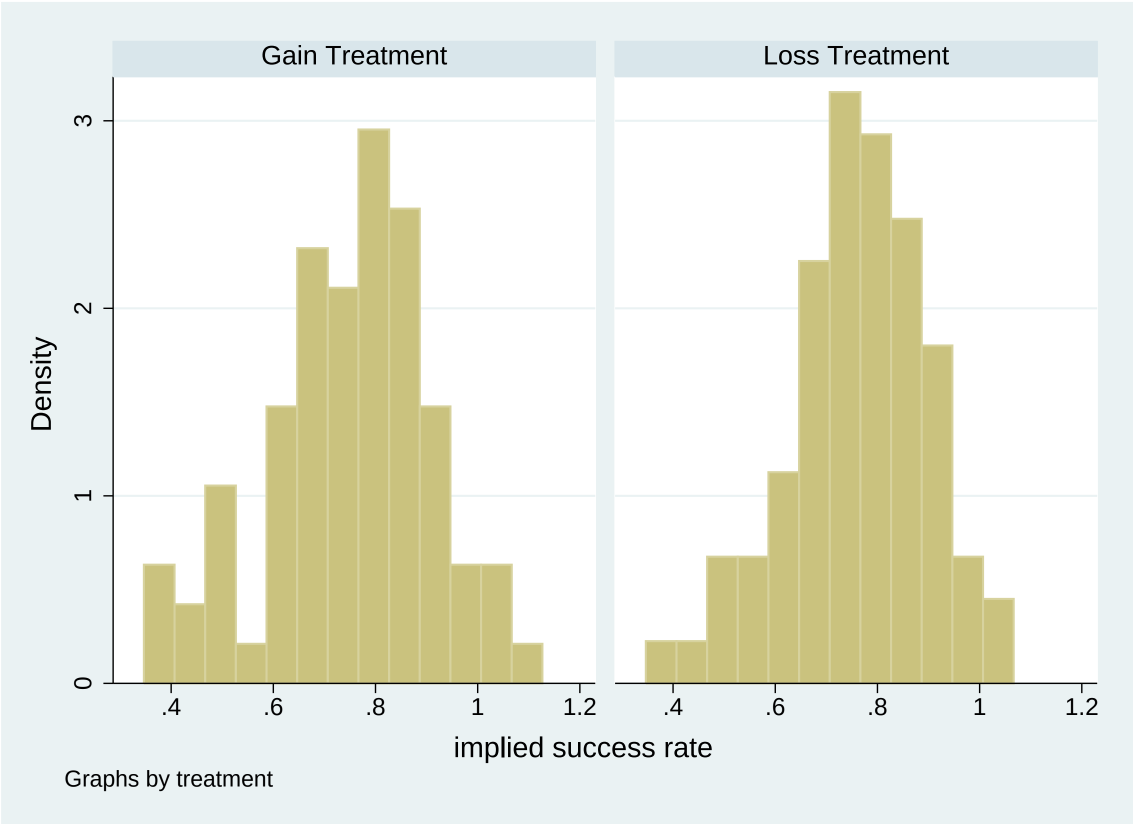

Figure 2: Predicted success rate, by treatment

admissible [0,1] interval. This suggests that large majority of subjects understood the scoring rules and put at least minimal effort in predicting their outcome. Thirdly, the picture provides us with an additional hint regarding subjects' minimum probability of success necessary to take the risk. Consider a student with implied success rate of .8. Presumably, she is quite sure about some of her answers. Her perceived confidence in the answer she is *least* certain about must be substantially below .8. This confirms our claim that the threshold probability is below .7 in most subjects.

It is interesting to consider the relationship between the predicted score and the number of missing answers. It is represented in Figure 3 which also shows a lowess approximation. The observed overall inversed U-shaped relationship is not surprising – with a large number of missing answers both very low and very high scores are excluded or extremely unlikely. It is more instructive to take a closer look at the number of missing answers for students with intermediate predictions. If our concern that students only cared about the grade (especially passing vs. failing) and thought they could predict their score precisely was valid, we would observe very little risk taking (a sudden hike in the number of missing answers) just above the announced threshold of 45 points. No such pattern can be seen.

Concerning our auxiliary hypothesis of males being more self-assure and/or more risk-seeking, gender effect was indeed significant – on average females left 8.04 questions unanswered while males only 6.86 ($p = .05$ in a MMW test). This was not associated with any superior knowledge – scores obtained (for either part of the exam) were nearly identical across genders.

By means of a multiple regression we may look for an interaction between gender and treatment. Table 2, where we also control for the grade obtained at tutorials, shows that there was no such effect.

Table 2: *Factors affecting the number of missing answers*

missing	Coef.	$P > \|t\|$
male	-1.252	0.153
male*LT	.716	0.570
LT	-.868	0.336
tut. grade	-1.815	0.000
cons	15.117	0.000

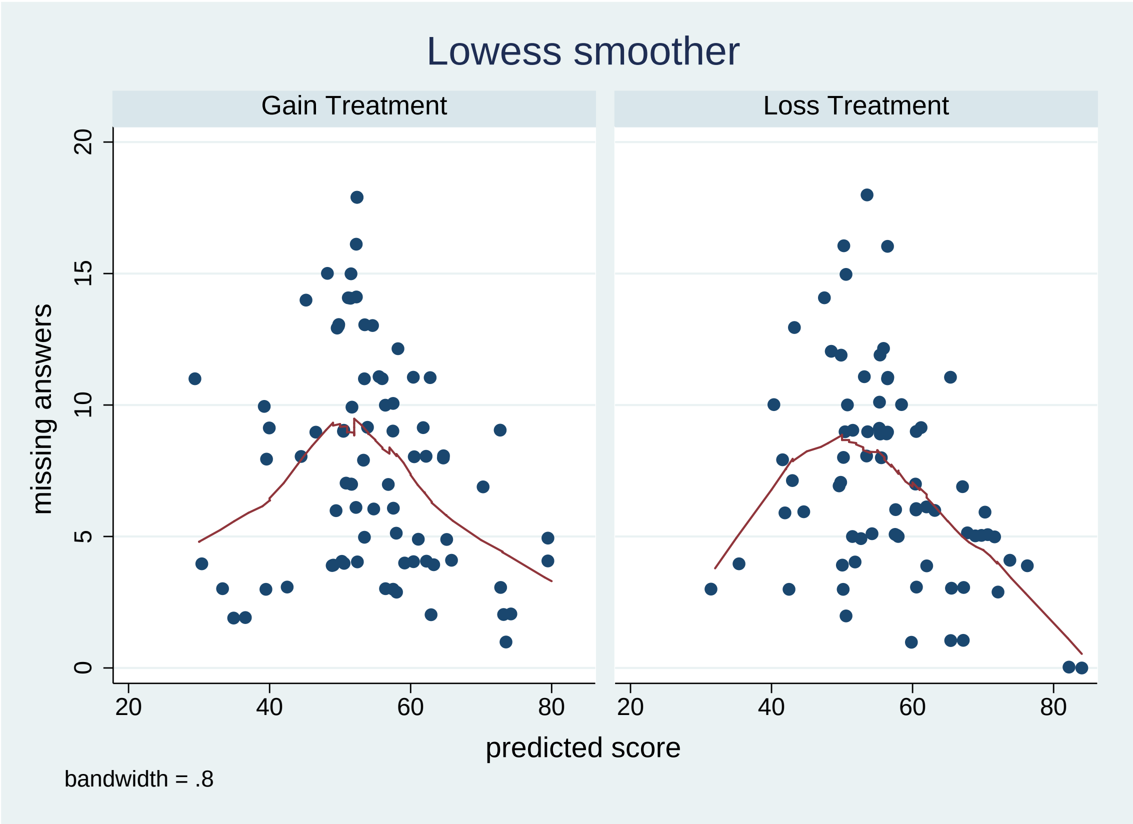

Figure 3: Predicted score and missing answers, by treatment

5 Conclusion

The current study was planned as a straightforward field test of the reflection effect. Despite a fairly large sample of subjects motivated to make best choices possible, such a pattern was hardly detected. This null effect cannot be ascribed to the specificity of the sample, in the sense that laboratory experiments on risk posture use university students as well. Nor was it likely that subjects did not notice or understand the scoring rules. Apparent deviation from the pattern predicted by Kahneman and Tversky could result from the fact that the underlying reward medium involved academic success rather than money; but to the extent that biases such as the reflection effect go away with experience, the all–too-familiar monetary domain is the one where we would least expect it to show. The study therefore casts some doubts on the empirical validity of the hypothesis of risk-seeking in losses (and thus the reflection effect).

References

Bosch-Domènech, A., and J. Silvestre (2006): "Reflections on gains and losses: A 2× 2× 7 experiment," *Journal of Risk and Uncertainty*, 33(3), 217–235.

Byrnes, J., D. Miller, and W. Schafer (1999): "Gender differences in risk taking: A meta-analysis.," *Psychological bulletin*, 125(3), 367.

Clotfelter, C. (1983): "Tax evasion and tax rates: An analysis of individual returns," *The review of economics and statistics*, 65(3), 363–373.

Kuhberger, A. (1998): "The Influence of Framing on Risky Decisions: A Meta-analysis* 1," *Organizational Behavior and Human Decision Processes*, 75(1), 23–55.

Kuhberger, A., M. Schulte-Mecklenbeck, and J. Perner (2002): "Framing decisions: Hypothetical and real," *Organizational Behavior and Human Decision Processes*, 89(2), 1162–1175.

Odean, T. (1998): "Are investors reluctant to realize their losses?," *The Journal of Finance*, 53(5), 1775–1798.

Thaler, R., and E. Johnson (1990): "Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice," *Management Science*, 36(6), 643–660.

TRAUTMANN, S., F. VIEIDER, AND P. WAKKER (2008): "Causes of ambiguity aversion: Known versus unknown preferences," *Journal of Risk and Uncertainty*, 36(3), 225–243.

TVERSKY, A., AND D. KAHNEMAN (1981): "The framing of decisions and the psychology of choice," *Science*, 211(4481), 453.

——— (1992): "Advances in prospect theory: Cumulative representation of uncertainty," *Journal of Risk and uncertainty*, 5(4), 297–323.

Faculty of Economic Sciences
University of Warsaw
44/50 Długa St.
00-241 Warsaw
www.wne.uw.edu.pl